

A COMPARISON OF AUDITORY AND VISUAL STIMULI IN DELAYED  
MATCHING TO SAMPLE WITH ADULT HUMANS

Anthony L. DeFulio, B.S.

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

December 2002

APPROVED:

Manish Vaidya, Major Professor  
Sigrid Glenn, Committee Member and Chair  
of the Department of Behavior Analysis  
Richard Smith, Committee Member  
David W. Hartman, Dean of the School of  
Community Service  
C. Neal Tate, Dean of the Robert B.  
Toulouse School of Graduate Studies

DeFulio, Anthony L., A comparison of auditory and visual stimuli in a delayed matching to sample procedure with adult humans. Master of Science (Behavior Analysis) December 2002, 57 pp., 2 tables, 14 illustrations, references, 19 titles.

Five humans were exposed to a matching to sample task in which the delay (range = 0 to 32 seconds) between sample stimulus offset and comparison onset was manipulated across conditions. Auditory stimuli (1" tone) and arbitrary symbols served as sample stimuli for three (S1, S2, S3) and two (S4 and S5) subjects, respectively. Uppercase English letters (S, M, and N) served as comparison stimuli for all subjects. Results show small but systematic effects of the retention interval on accuracy and latency to selection of comparison stimuli. The results fail to show a difference between subjects exposed to auditory and visual sample stimuli. Some reasons for the failure to note a difference are discussed.

## ACKNOWLEDGMENTS

I would like to thank the entire faculty of the Department of Behavior Analysis for their assistance both on this project and over the course of my years here at the University of North Texas. You have all been wonderful.

I would like to extend extra thanks to the members of my committee, Rick Smith, Sigrid Glenn and Manish Vaidya, who have each played a critical role in my development as a behaviorist, as a scientist, and as a human being.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS .....	ii
LIST OF TABLES.....	iv
LIST OF ILLUSTRATIONS.....	v
Chapter	
1. INTRODUCTION .....	1
2. GENERAL METHOD.....	7
3. TRAINING .....	14
4. DELAYED MATCHING TO SAMPLE .....	21
5. DISTRACTION.....	42
6. DISCUSSION.....	49
REFERENCE LIST .....	56

## LIST OF TABLES

Table	Page
1. Total number of sessions, number of sessions per condition, and number of trials in each session.....	13
2. Error analysis for distraction conditions .....	48

## LIST OF ILLUSTRATIONS

Figure	Page
1. Sample stimuli, comparison stimuli, and their relations for each subject .....	9
2. Training Accuracy .....	16
3. Training Latency .....	19
4. DMTS Accuracy (S1, S2, S3).....	23
5. DMTS Accuracy (S4, S5).....	25
6. DMTS Latency (S1, S2, S3) .....	27
7. DMTS Latency (S4, S5) .....	29
8. Latency to comparison selection as a function of retention interval with regression lines fit to individual and pooled data.....	32
9. S1 error analysis.....	34
10. S2 error analysis.....	36
11. S3 error analysis.....	37
12. S4 error analysis.....	39
13. S5 error analysis.....	40
14. Distraction Accuracy .....	45

## CHAPTER 1

### INTRODUCTION

Remembering as a phenomenon has been the subject of inquiry since the inception of experimental psychology. Wundt started the first psychological laboratory at approximately the same time Ebbinghaus was conducting the first experimental investigations of memory and forgetting (Schneider, 2000). Over the last thirty years the inquiry concerning remembering has been guided by at least two basic approaches.

The more common approach involves investigations of the structures and internal processes that may mediate remembering. An alternative approach is to investigate the act of remembering itself. From the latter perspective, remembering has been defined as behavior under discriminative control of a stimulus where “reinforcement is ... made contingent on appropriate behavior in the absence of the stimulus” (Palmer, 1991).

Psychologists have studied the phenomenon of remembering using a wide variety of procedures including free recall, cued recall, and recognition tasks (Roediger, Marsh and Lee, 2002). These tasks require subjects to, for example, say yes or no in the presence of stimuli that may or may not have been a member of the original sample-set at some point after the original set was presented (Roediger, Marsh and Lee, 2002).

Another important preparation is the delayed-matching-to-sample (DMTS) procedure (Blough, 1959) in which subjects are required to engage in conditional discriminations in the absence of the sample stimuli. A typical trial in Blough’s (1959)

experiment began with the presentation of a sample stimulus for one second. A sample stimulus was either a flickering or steady light illuminated from an aperture between two response keys. At some point after the offset of the sample stimulus, the response keys were illuminated with a flickering light and a steady light, respectively. The pigeon made a choice response by pecking one of the keys. If the response was on the key displaying the stimulus that matched the sample stimulus then a reinforcer was delivered, followed by an inter-trial interval (ITI). A response to the key displaying the non-matching stimulus simply turned off all lights and began the ITI.

The period of time between the offset of the sample stimulus and the onset of the illumination of the comparison keys changed from trial to trial and constituted the main independent variable. Blough (1959) found that accuracy decreased as delay between sample offset and comparison onset (hereafter, retention interval) increased. This finding has been replicated many times in studies investigating non-human remembering (Hunt, Parr and Smith, 1999; Jones and White, 1992) and is, in fact, the most consistent finding in studies of non-human remembering that use the DMTS procedure (Sargisson and White, 1999).

In addition, a host of other factors have been shown to reliably affect performance in DMTS tasks with non-humans. For example, changes in reinforcement ratio have been shown to produce bias toward the stimulus associated with the greater ratio of reinforcement (Jones and White, 1992), an effect that becomes more pronounced with increases in retention interval. Changes in the ratio of presentation of each of the sample stimuli within a session have had similar effects (Parr, Hunt, and Williams, 1999).



However, the bias toward the more frequently presented sample stimulus found in the latter study did not change as a function of changes in retention interval.

Other manipulations involving the sample stimuli have resulted in equally orderly effects on DMTS performance. White (1985) used colored keys as sample stimuli to study performance of pigeons in a DMTS procedure. White's data show that as the difference between the wavelengths of the sample stimuli is increased performance on the task improves. This manipulation changes the Y-intercept but not the slope of the function relating accuracy to delay, indicating that this change is a function of increased initial discriminability.

A factor that repeatedly has been shown to have a strong influence on performance is the duration of the sample stimuli (Nelson and Wasserman, 1978; White, 1985; Wixted and Ebbesen, 1991). Nelson and Wasserman (1978) showed that as the duration of the sample stimulus was increased, accuracy on the DMTS task increased across a series of retention intervals, although the effect was more pronounced at shorter retention intervals. Other methods of increasing exposure to the sample stimuli have had the same result. White (1985) re-analyzed data from both Roberts (1972) and Grant (1981) and showed this effect with increases in the response requirement for observing or with increases in the number of times the sample stimulus was presented before a trial progressed to the choice phase.

In sum, there is considerable agreement about the conditions under which the accuracy of non-human performance decays as a function of the value of the retention interval. As mentioned above, experiments with non-human subjects show that changes in the retention interval do not consistently alter the effects of initial discriminability or

imposed biases. Retention intervals have been shown to affect accuracy of performance however, and this effect is modulated by the time the subject spends in the presence of sample stimuli.

The same cannot be said for humans. Although studies have shown effects similar to those demonstrated in studies of non-human performance, the effect has been limited to certain populations under certain conditions. For example, Parsons, Taylor, and Joyce (1981) found that increases in retention interval decreased typically developing five-year-old children's accuracy in a DMTS task that required subjects to emit an observing response to begin the delay interval and a second response (effective only after a fixed-interval has passed) to produce the comparison stimuli. However, this effect was largely eliminated for children trained to perform differential collateral responses in the presence of the sample stimuli, a finding later replicated by Torgrud and Holborn (1989). Both of these findings were interpreted in terms of response chains. The critical factor that determined the effect of increasing retention interval on performance accuracy was whether different precurrent operant responses were established for each sample stimulus.

Similar to studies involving non-humans, studies of human remembering have implicated aspects of the sample stimuli as critical determinants of performance in DMTS tasks. For example, neurological patients were more likely to make errors to comparison stimuli that were similar to sample stimuli along a common stimulus dimension than stimuli that were markedly different along that dimension (Sidman, 1969). In a later study, Baron and Menich (1981) speculated that increasing the number of elements in the sample stimuli might decrease accuracy, but found that although latency to comparison

selection increased with increasing retention intervals, accuracy was relatively stable.

The consistencies in the non-human literature are not maintained in the context of arbitrary matching procedures. For example, studies of DMTS performance with pigeons report slower acquisition of the conditional relations and lower overall accuracy with auditory sample stimuli than with visual sample stimuli (Kraemer and Roberts, 1984). In contrast, dolphins have shown high levels of accuracy in DMTS tasks with auditory stimuli as samples at retention intervals considerably longer than delays typically used with pigeons (Herman and Gordon, 1974).

Dube et al. (1989) suggest that pigeons' performance on typical DMTS procedures where sample and comparison stimuli are identical is qualitatively different from the performance of humans on such tasks because of the ease with which humans display generalized identity matching and, unlike humans, pigeons in DMTS tasks exhibit explicitly trained, sample-specific behaviors. Dube's suggestion implies that a DMTS task for humans consisting of arbitrary conditional discriminations would yield results more comparable to the non-human literature. Auditory-to-visual conditional discriminations are arbitrary by nature, and therefore may be advantageous for use in preparations designed to study human remembering.

Studies investigating remembering in humans have typically shown less decrement in performance when auditory samples are used than when visual samples are used (Kraemer and Roberts, 1984). However, the stimuli used in these procedures are real words and the procedures include free recall, serial recall, and recognition tests, which, as described in the opening paragraph, are substantially different from delayed matching to sample tasks. The use of real words reduces the usefulness of comparisons

between findings from this body of literature and data from non-human DMTS performance, especially in light of Dube et al.'s (1989) suggestion.

The present experiment asks several questions. First, do subjects exposed to auditory stimuli perform as accurately on a DMTS task as those exposed to visual stimuli? Also of interest is whether changes in retention interval affect accuracy differently for these subjects. In addition to accuracy, questions concerning latency to respond are also investigated. Specifically, will changes in retention interval have systematic effects on either latency to observing response or latency to comparison selection, and are these effects dependent on stimulus modality? These questions were addressed by comparing the performances of humans exposed to auditory versus visual stimuli as samples in a DMTS task across increasing retention intervals.

## CHAPTER 2

### GENERAL METHOD

#### Subjects

One female (S1) and four male undergraduate students from the University of North Texas, ranging from 18-27 years of age, participated in this experiment. The participants were recruited from introductory behavior analysis classes, newspaper advertisements, and flyers posted on the University campus. Participants earned \$1.00 for attendance each session and three cents of bonus pay for each correct response in a session. The earned pay (per hour) over the course of the experiment ranged from \$6.72 to \$10.69. Subjects were told that they would be paid in a lump sum at the end of the experiment and that terminating their participation before the end of the study would result in a forfeiture of their bonus earnings. The attendance fee, however, was theirs to keep.

#### Setting and Apparatus

Experimental sessions were conducted in a 6'x 4.7' room, which contained a table, the apparatus (described below), and two chairs. Participants worked alone (during most sessions) in the room with the door closed and were instructed not to bring any personal effects or other materials inside the room.

The apparatus consisted of an Apple G3® notebook computer (©Apple Computer Inc., [www.apple.com](http://www.apple.com)) fitted with a Troll Touch® TouchSTAR™ touch screen adapter (©Troll Touch/T2D Inc., [www.trolltouch.com](http://www.trolltouch.com)). Auditory stimuli were presented via

Sony® headphones, model MDR-009 (©Sony Corporation of America, [www.sony.com](http://www.sony.com)).

Visual stimuli and visual accompaniment of auditory stimuli were presented on the monitor behind the touch screen adapter. Subjects responded by touching the touch screen adaptor. A program (MTS v 11.6.7) written by Bill Dube and colleagues controlled all stimulus presentation, data collection, and management of contingencies.

### General Procedure

The experimental preparation consisted of a conditional discrimination task in which tones (for subjects S1, S2, and S3) or visual forms (for subjects S4 and S5) served as sample stimuli and lowercase English letters served as comparison stimuli. Subjects learned simultaneous conditional relations with a delayed-prompt procedure (described below). When the performance was accurate and stable (criteria described below) the task was changed such that observing responses removed sample stimuli in addition to producing the comparison stimuli. Delay between sample offset and comparison onset were manipulated across subsequent conditions.

For S1, S2, and S3, three pure tones (196.00, 277.18, and 369.99 Hz) were used as sample stimuli. Each tone was generated by a pure tone generator in SoundEdit™ (©Macromedia Inc., [www.macromedia.com](http://www.macromedia.com)) and presented via headphones. The presentation of the auditory stimulus was accompanied by a 1.5-inch black square in the center of the monitor. For S4 and S5, the sample stimuli consisted of arbitrary black figures that were 1.5 x 1.0 inch tall and bore no obvious resemblance to objects likely to be encountered by the subject in everyday life. These stimuli are shown in Figure 1. Sample stimuli were presented quasi-randomly with the provision that the same sample

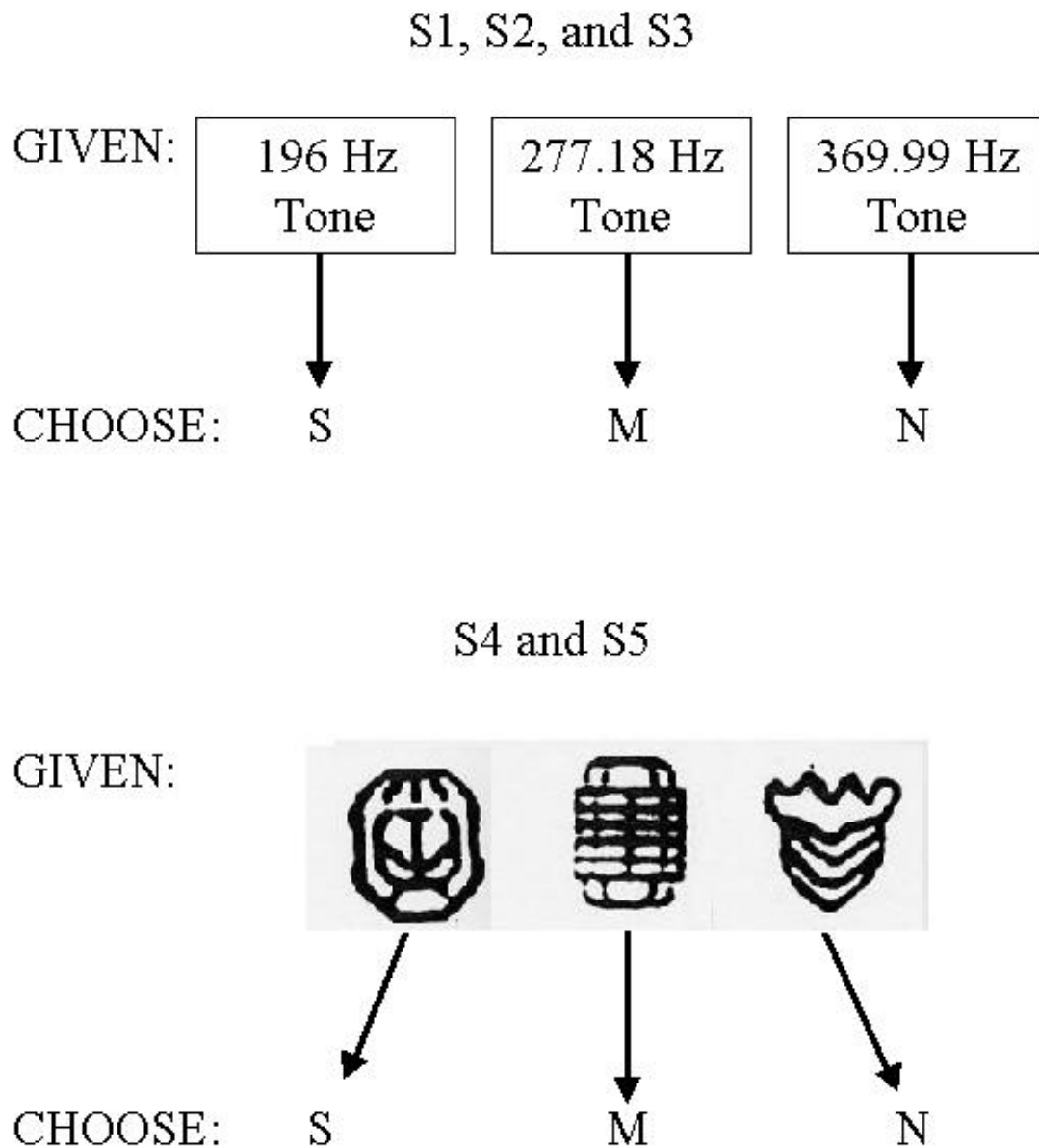


Figure 1. Sample stimuli, comparison stimuli, and their relations for each subject.

stimulus could not be presented consecutively for more than three trials. For all subjects, the uppercase English letters M, N, & S (24-point Geneva font in black) comprised the array of comparison stimuli. Each comparison stimulus could appear in any of the four corners of the screen. Positions of comparison stimuli were quasi-randomized across trials with two provisions: that the “correct” choice not appear in the same position for more than three consecutive trials and that a stimulus, regardless of its status during the trial, not appear in the same position for more than three trials.

The presentation of three comparison stimuli in any of four positions allows for 24 possible arrays. Three conditional relations were included in the task, yielding 72 possible trials in which all possible combinations of sample and comparison stimuli were presented. A single session consisted of either one or two exposures to all 72 possible combinations, for a total of either 72 or 144 trials. Table 2 shows the number of trials per session for each subject. Subjects S1, S2, and S4 began the experiment with a single exposure to each permutation, but sessions were increased to 144 trials to increase possible earnings in an effort to retain subjects. Sessions were later reduced to 72 trials for all participants when increases in the retention interval (see below) caused session lengths to increase beyond those specified in the informed consent forms.

Trials began with the presentation of one of three sample stimuli. Auditory sample stimuli were repeated every 1.5 seconds until subjects touched the black square (observing response). A lower limit was imposed on the latency to the observing response such that the auditory stimulus had to be presented once completely before observing responses were effective. An observing response removed the black square and produced an array of comparison stimuli. A touch to one of the comparison stimuli



(comparison selection) turned off the auditory stimulus and cleared all visual stimuli from the display. If the participant touched the experimenter-designated correct stimulus the word “CORRECT” appeared in the center of the screen accompanied by a brief sequence of two beeps. This was followed immediately by a 1.5 second inter-trial interval (ITI). The screen was entirely white during the ITI and any responses to the screen reset the ITI. Selection of the “incorrect” comparison stimulus cleared the display and began the ITI immediately. Trials with visual sample stimuli were identical with the exception that no lower limit was placed on the observing response.

Each participant was exposed to a brief session before the experiment began. The goal of this brief exposure was to familiarize the subjects with the procedures and to give them an opportunity to withdraw from the study after exposure to the task and the setting but before substantial time was invested. The session began with a participant seated in front of the computer. The experimenter then read the following instructions out loud:

To begin the session you will press the key labeled return on the keyboard. When the session begins make your responses by touching the screen. All responses should be made by touching the screen. When the session is over leave the room and let me know you are finished. Please put the headphones on before you begin.

The brief exposure consisted of 12 simultaneous conditional discrimination trials with delayed prompts. These trials were identical to training trials described below. Upon completing the session individuals were asked if they wished to continue participation in the study. Individuals who chose to participate waited approximately 10 minutes between the brief session and the first full session, which was conducted on the

same day. Subsequent to the first session no more than one session was conducted per day for an individual participant. Sessions were conducted five days per week. Total number of sessions, as well as the number of sessions per condition, are presented in Table 1.

	Train	0	1	2	4	8	16	32	64	GIM	SAM	COM	REA	TOTAL
Subject														
S1	4 <sup>a</sup>	1 <sup>a</sup> /4	5	4	3	6	4/5 <sup>a</sup>	-	-	-	2 <sup>a</sup>	-	3 <sup>a</sup>	41
S2	1 <sup>a</sup> /7 <sup>b</sup>	3	3	3	3	5	3/1 <sup>a</sup>	3 <sup>a</sup>	-	-	3 <sup>a</sup>	-	4 <sup>a</sup>	39
S3	3	3	3	3	3	3	3/1 <sup>a</sup>	3	-	-	2 <sup>a</sup>	-	3 <sup>a</sup>	30
											1 <sup>a,d</sup>			
S4	4 <sup>a</sup> /1	3	3	3	3	3	3	2/1 <sup>a</sup>	3 <sup>a</sup>	3 <sup>a</sup>	5 <sup>a</sup>	2 <sup>a</sup>	3 <sup>a</sup>	45
								3 <sup>a,c</sup>						
S5	3	3	3	3	3	3	4/1 <sup>a</sup>	5 <sup>a</sup>	5 <sup>a</sup>	-	2 <sup>a</sup>	-	4 <sup>a</sup>	41
								2 <sup>a,c</sup>						

a = 72 trial session(s)

b = session five was unique, see text for details

c = second exposure to the 32-second retention interval condition  
occurred after the 64-second retention interval condition

d = second exposure to the SAM condition occurred after the REA condition

Table 1. Total session, sessions per condition, and trials per session for each subject. Note that when both 72-trial and 144-trial sessions were conducted in the same condition that these sessions are separated within a cell. Conditions are presented chronologically left to right except as noted.

## CHAPTER 3

### TRAINING

#### Procedure

All subjects were first taught three simultaneous conditional relations using a delayed prompting procedure in which the distracter comparison stimuli were removed following a period during which no responses occurred. This period, hereafter referred to as the delay-to-prompt, was always set at 0.1 seconds at the beginning of a session. The delay-to-prompt was increased by one-third of its previous value following trials with correct responses. A correct response was defined as touching the experimenter-designated “correct” comparison stimulus. After the delay-to-prompt was greater than two seconds it increased by 0.5 second following each correct response. Incorrect responses resulted in a reduction of the delay-to-prompt in similar fashion. Touching the  $S^-$  stimuli or the location where the  $S^-$  stimuli had been displayed prior to the prompt were counted as errors.

Trials began with the presentation of a sample stimulus (see above). For S1, S2, and S3, touching the black square (hereafter, observing response) removed the black square from the display and produced the array of comparison stimuli. The tone continued to play in the manner described above. For subjects S4 and S5, touching the sample stimulus added the array of comparison stimuli to the display. Consequences for accurate and inaccurate responses, whether prompted or unprompted, were identical to

those in the 12-trial session (described above). The training condition ended when participants completed a minimum of three sessions in which at least 80% of trials were unprompted and a minimum of two sessions in which at least 90% of unprompted trials were correct. The only exception to this was subject S2, who relied entirely on the prompt for the first three sessions of the experiment. For this reason, S2 was exposed to a no-prompt successive conditional discrimination during the fourth session, before which the following instruction was delivered:

Your behavior determines the length of the session.

After the fourth session S2 returned to the training sessions until his responses met the criteria for ending the training condition.

### Results

Figure 2 shows the percent of trials that were unprompted in each training session (open circles), as well as the percent of unprompted trials with a correct response (filled circles). The left column shows data from three subjects exposed to audio-visual conditional discriminations, and the right column shows data from two subjects exposed to visual-visual conditional discriminations.

This figure shows that S1 learned the conditional relations to a high degree of accuracy during the first session. S1's performance was greater than 90% accurate across subsequent training sessions. The proportion of trials without a prompt also remained consistently high. In fact, S1 only experienced the prompting when the programmed delay-to-prompt was shorter than her typical latency to select a comparison stimulus. After the first session no prompts were contacted or experienced when the delay-to-

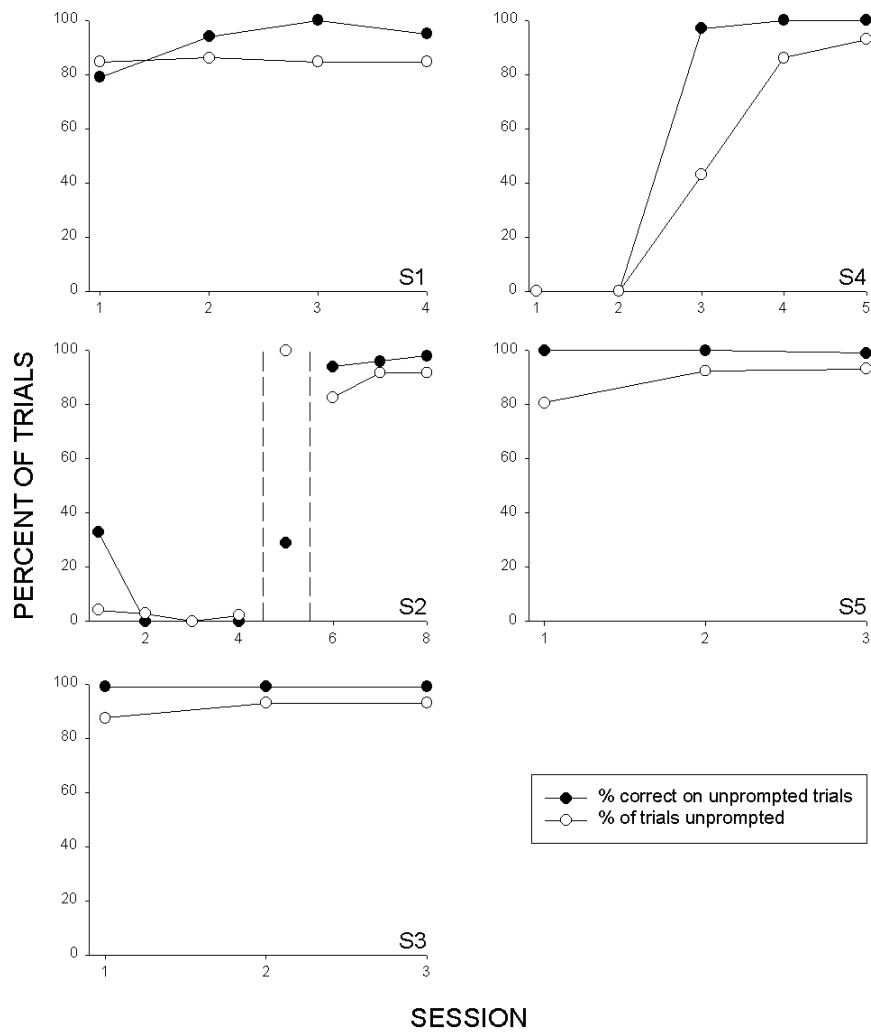


Figure 2. Percentage of unprompted trials in which subjects selected the correct comparison stimulus and percentage of trials in which subjects responded before the prompt during training sessions.

prompt exceeded 1.1 seconds. The graph for S3 (lower left) shows that S3's performance could be characterized in a similar manner. The proportion of unprompted trials and the accuracy of comparison selection on those trials were high from the beginning of training and remained high throughout the training condition.

Subject S2's performance was different in that he waited for the prompt on nearly every trial during the first four sessions. After four sessions of almost completely prompt-dependent performance, S2 was exposed to a single session in which the conditional discrimination task was changed from simultaneous to successive and no prompts were delivered. In addition, the subject was told that his behavior determined the length of the sessions. Accuracy was at approximately chance levels during this session (29% correct). After this session S2's performance became similar to other subjects' performances, with highly accurate performance and relatively little use of prompts.

The top right graph shows that S4 waited for the prompt from the outset through most of the third session. S4 began responding before the prompt consistently during the latter half of the third session. His performance for the remainder of the training condition was similar to other subjects, including both a high degree of accuracy and a high proportion of unprompted trials.

The graph in the middle of the right column presents data from S5. This graph shows that S5's performance was highly accurate from the beginning of the training condition. This graph also shows that the proportion of unprompted trials was high from the outset and remained high throughout the training condition.

In sum, all subjects' performance on the conditional relations was highly accurate by the end of training. Three of the five subjects acquired the conditional relations rapidly with minimal use of prompts. Two of the five subjects initially waited for the prompt on nearly every trial. S2 was exposed to an instruction and a change in contingency for one session, after which S2's performance was similar to that of other subjects. S4's prompt dependence changed without changes in contingencies. Both subjects' performance resembled the performance of other subjects after prompt dependency ceased.

Figure 3 shows the median latencies for observing responses (filled circles) and the selection of comparison stimuli (open circles) across sessions. Data for subjects exposed to audio-visual conditional discriminations (S1, S2, and S3) are in the left column, and data for subjects exposed to visual-visual conditional discriminations (S4 and S5) are in the right column.

The data show that the latency to observing response was generally below one second. Subjects S1, S2 and S3's latencies were stable at about one second with no increasing or decreasing trends across sessions. Subjects S4 and S5's observing response latencies were shorter. For example, S4's median latencies ranged from 0.56 to 0.7 seconds, with no increasing or decreasing trend across sessions. In the beginning of training S5 showed a median latency of over one second, but latency to observing response decreased steadily for this subject over the course of training sessions, reaching 0.64 seconds for the third session.

Graphs in the left column show that subjects S1, S2, and S3's performances were similar to one another in terms of latency to comparison selection. Median latency to



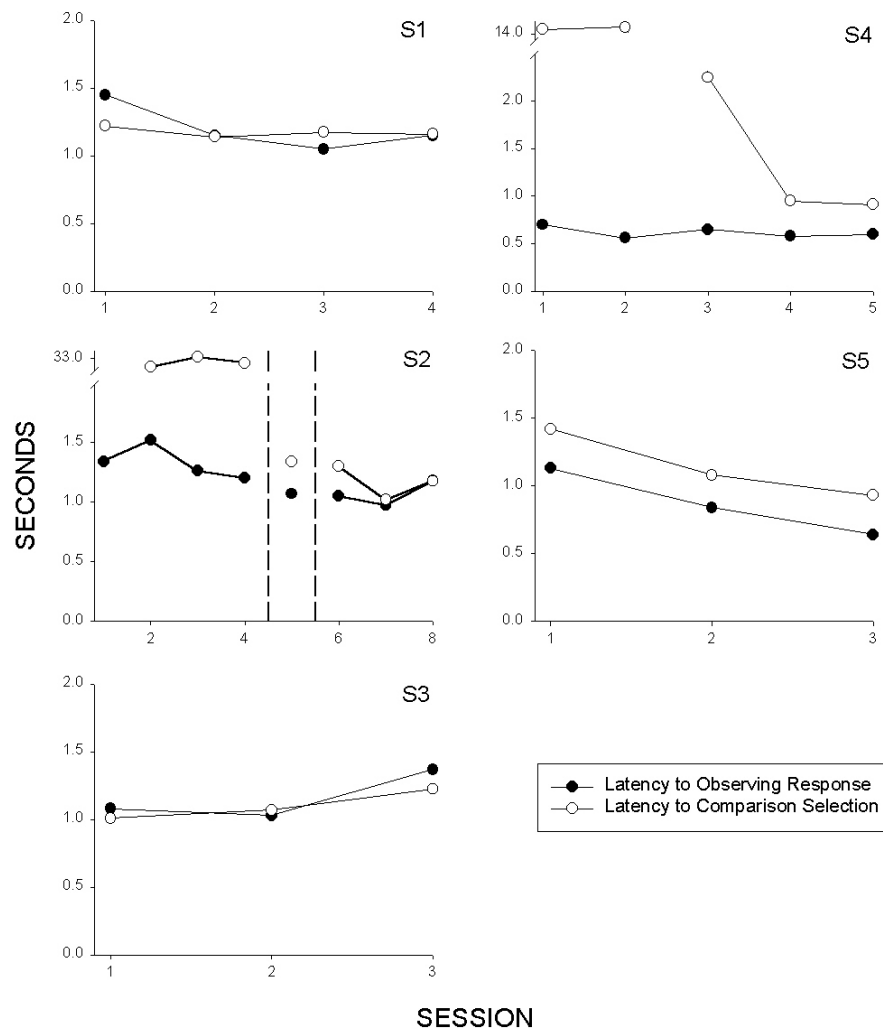


Figure 3. Median latency per session to observing response and to comparison selection.

select comparison stimuli was consistently below 1.5 seconds. The only exception to this is S2's performance in the first few sessions (described above) in which he waited for prompts on every trial. S4's performance was similar to that of subject S2, showing initially long latencies to comparison selection due to the tendency to wait for the prompt on most trials. As with accuracy of performance, the latency to select comparison stimuli became similar to those of other subjects when he began to respond before the prompt. Subject S5 also showed consistently short latencies to touch comparison stimuli. Similar to subjects S1 and S3, subject S5's performance did not change significantly over the course of the training condition in terms of median latency to select comparison stimuli.

In sum, median latency to observing response measures were typically slightly greater than one second for subjects exposed to tones as sample stimuli and roughly 0.6 seconds for subjects exposed to arbitrary visual forms as sample stimuli. This is an artifact of the procedural requirement that no observing response could register until the auditory sample stimulus had been presented completely at least once. Median latency to comparison selection typically approximated one second for all subjects across training sessions. Subjects S4 and S2 showed initially long latencies to comparison selection, which was directly related to their reliance on the prompt. After these subjects began to respond before the prompt their latency to comparison selection became consistent with those of the other three subjects.

## CHAPTER 4

### DELAYED MATCHING-TO-SAMPLE

#### Procedure

Following acquisition of the three conditional relations, the simultaneous matching-to-sample procedure was changed to a successive matching-to-sample procedure with a zero-second delay between sample offset and comparison onset. Subjects were then exposed to a series of conditions in which the delay between the offset of sample stimuli and the onset of comparison stimuli (hereafter, retention interval or RI) was manipulated. Retention interval values were held constant within a condition and changed across conditions. All subjects began this phase with the RI set at zero seconds. Across conditions, subjects were exposed to an ascending series of RI values of 1, 2, 4, 8, 16, and 32 seconds. Conditions were changed when a subject's performance was deemed stable. The criteria for stability included at least three sessions in a condition with no increasing or decreasing trend in the accuracy of performance. Table 2 shows RI values and the number of sessions in each condition for each subject. S4 and S5 were exposed to an additional condition (RI = 64 seconds), and then returned to the 32-second condition before moving to other conditions.

There were a few errors in the training and testing procedures described above. Subjects S4 and S5 were both exposed to a single session in which auditory sample stimuli were used instead of the typical two-dimensional arbitrary forms due to

experimenter error. For S4 this occurred between the RI = 8 second and RI = 16 second conditions, and for S5 this occurred after a single session in the RI = 16 second condition. Subject S1 experienced two sessions in the RI = 0 condition (session numbers 6 and 7) in which the auditory sample stimulus failed to play on some trials. This was due to a programming error. Data from these sessions are not included in the following analysis.

## Results

Figure 4 presents the proportion of trials with a correct response across sessions for S1, S2, and S3 in the top, middle, and bottom graphs, respectively. The top graph (S1) shows highly accurate performance during the zero-second retention interval. In all subsequent conditions, S1's performance started out accurate, dropped to a lower level of accuracy, and then became more accurate again. This pattern of accuracy decline followed by increased accuracy is seen across all retention intervals for this subject. Increases in accuracy did not occur as quickly when retention interval was eight or 16 seconds. Accuracy was lowest during the 16-second retention interval, but never dropped below 85%.

The middle graph presents percent of correct trials for S2. This graph shows that S2's performance remained highly accurate across conditions. There were no systematic changes in percent of correct trials across retention intervals for this subject. Within the two, four, eight, and 16-second retention interval conditions, the lowest accuracy during the condition was during the first session. The bottom graph presents percent of correct trials for S3. This graph shows that S3's performance was highly accurate across all retention intervals, with a minimum of 97% correct across these conditions. As with subject S2 there were no systematic changes in accuracy across retention intervals

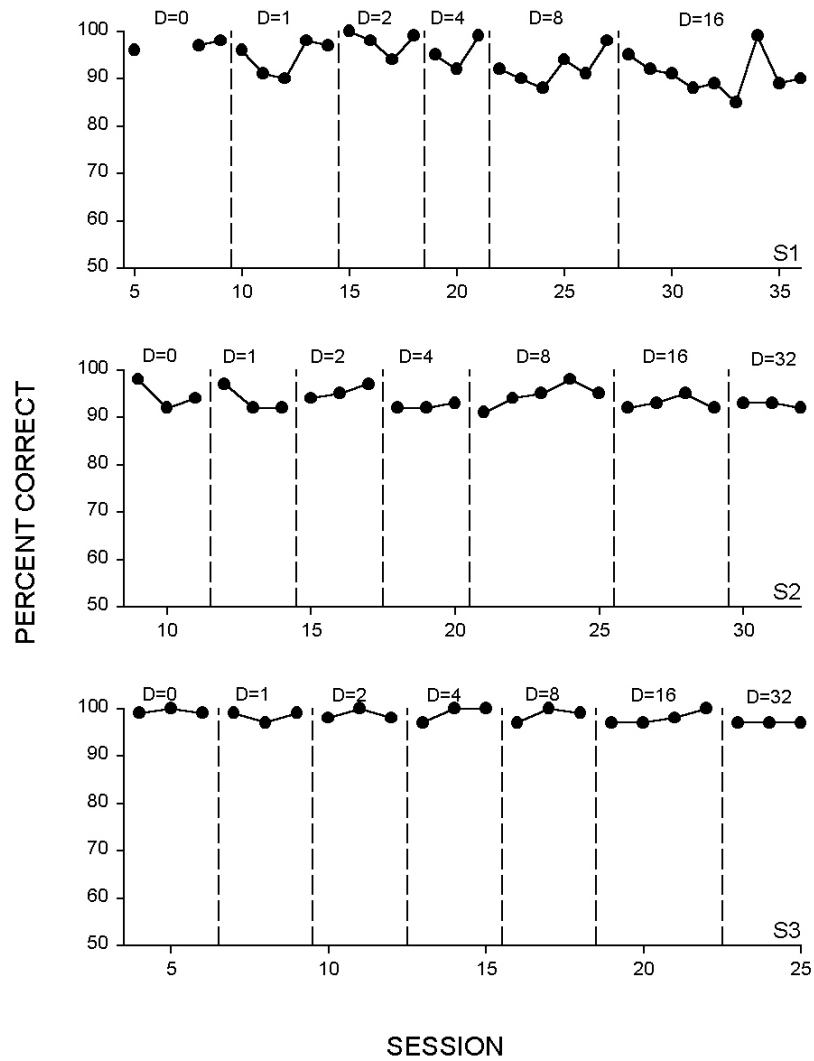


Figure 4. Percentage of correct responses across retention interval conditions for subjects exposed to tones as sample stimuli. Retention interval values are indicated above each graph.

for this subject, however accuracy in the 32-second retention interval condition never exceeded 97%. Also similar to S2, within the two, four, eight, and 16-second retention interval conditions the least accurate performance was observed during the first session in a condition.

Figure 5 presents percent of correct trials in the delayed-matching-to-sample conditions for subjects presented with arbitrary forms as sample stimuli. The top graph shows that S4's performance was highly accurate from the outset of the delayed-matching-to-sample conditions. S4's performance remained at least 99% accurate until the first exposure to the 16-second retention interval. Performance dropped to the lowest point for a single session (92%) during the 32-second condition.

The bottom graph shows that S5's performance was highly accurate through the 8-second retention interval condition. Percent of correct trials never drops below 94% during this period. Accuracy measures in some conditions (2, 4, and 64-second RI) followed a pattern in which lowest accuracy within a condition occurred during the first exposure to that condition. Overall accuracy decreased as the retention interval increased. Additionally, the first exposures to each condition showed a pattern of decreasing accuracy across conditions systematically beginning with the 8-second retention interval condition. Within the 16 and 32-second retention interval conditions S5 showed a decrease in accuracy after the first exposure to that condition, followed by a recovery. This is similar to the within condition pattern of accuracy commonly exhibited by S1.

In sum, increases in the retention interval value had no effect on the accuracy of performance for two subjects (S2 and S3), a very slight effect for one subject (S4) and a

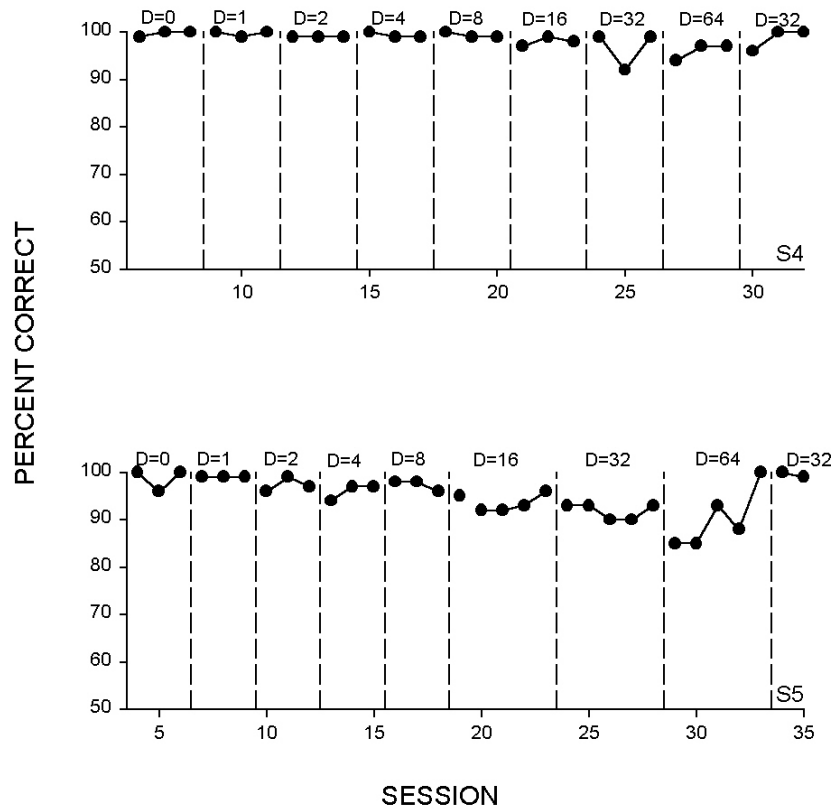


Figure 5. Percentage of correct responses across retention interval conditions for subjects exposed to arbitrary forms as sample stimuli. Retention interval values are indicated above each graph.

more noticeable effect for two subjects (S1 and S5). Accuracy remained high for all subjects throughout all retention interval conditions and did not drop below 90% for any subject during any condition except the 64-second retention interval condition for subject S5, in which accuracy was 85%. Within a condition, the data show that two patterns predominated. These patterns were not mutually exclusive. One pattern is characterized by the lowest percent of correct trials within a condition occurring on the first exposure to that condition. This pattern is present in data for subjects S2 (four of seven conditions), S3 (four of seven conditions), and S5 (three of eight conditions). The second pattern includes a decrease in accuracy after the first exposure to a condition followed by a recovery. Figures 4 and 5 show that subjects S1 (five of six conditions) and S5 (two of eight conditions) both exhibited this pattern. Subject S4 showed no within condition pattern of changes in accuracy.

Figure 6 shows median latencies across the conditions in which RI values were manipulated for subjects S1, S2, and S3 (top, middle and bottom graphs, respectively). Filled circles represent latency to the observing response and open circles represent latency to comparison selection.

Graphs for S1, S2 and S3 show relatively consistent latency to produce the comparison array across conditions. Median latency to observing response ranged from 0.97 seconds to 1.47 seconds for these three subjects. S2 showed a slight increase in latency to observing response as retention interval value increases. Median latency to observing response did not change as a function of changes in retention interval value for subjects S1 and S3.



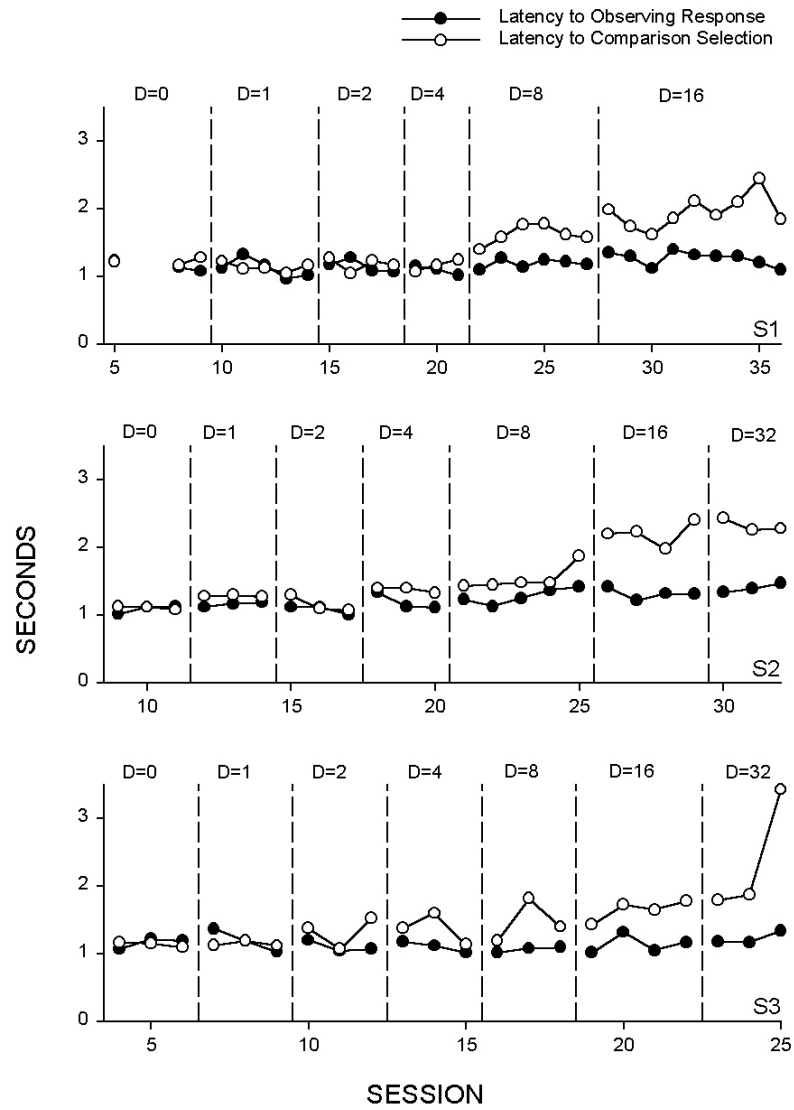


Figure 6. Median latencies across retention interval conditions for subjects exposed to tones as sample stimuli. Retention interval values are indicated above each graph.

Median latencies to select comparison stimuli remained short for the first few conditions and then began to increase as RI values increased. Latency to comparison selection became increasingly divergent from latency to observing response between the 4-second and 16-second retention interval conditions for S1 and S2, and between the 2-second and 32-second retention interval conditions for S3. For all three subjects median latency to comparison selection values increased as retention interval increased.

Figure 7 shows median latencies subjects S4, and S5. As with the previous figure, filled circles represent latency to observing response and open circles represent latency to comparison selection.

The graph for S4 (top) shows that latency to observing response did not change through the 8-second retention interval condition. Latencies were brief during this period, with median values typically around 0.6 seconds. Beginning with the 16-second condition, observing response latencies increased slightly, with median values typically around 0.8 seconds.

The top graph also shows that S4's latency to select comparison stimuli. This figure shows that latency to comparison selection was typically about one second during the first three conditions. Beginning with the RI = 8-second condition latencies to comparison selection began to increase. This effect became more pronounced when the retention interval was increased to 16 seconds. Latency to comparison selection lengthened again when retention interval was increased to 32 seconds, but median latencies during the 64-second retention interval condition all fall within the range of median latencies from 32-second retention interval conditions.

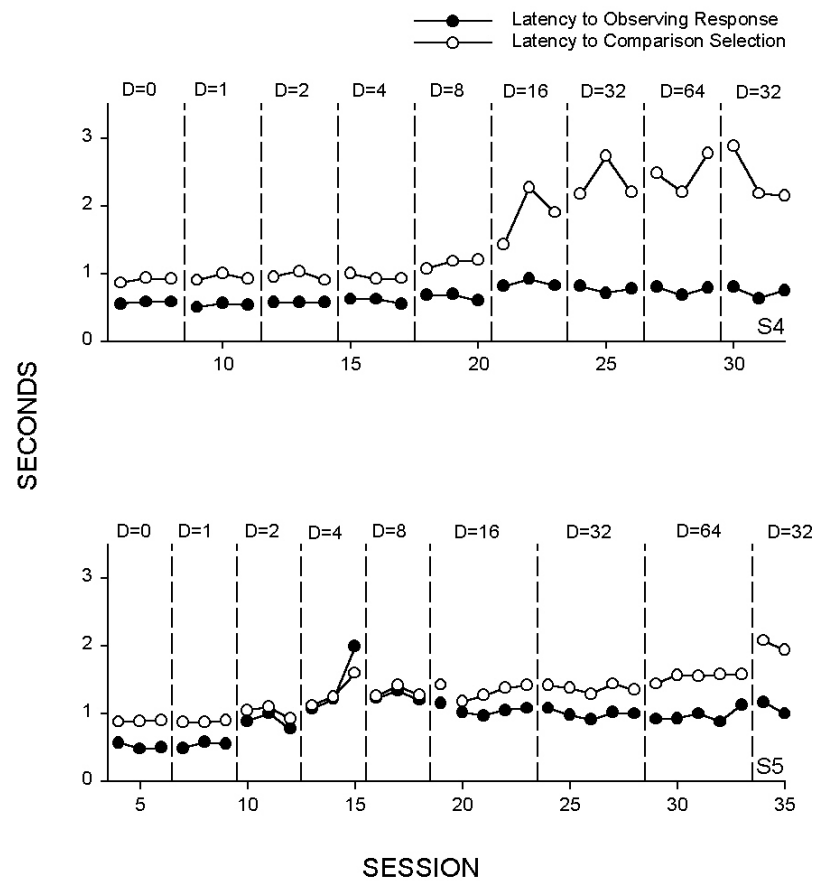


Figure 7. Median latencies across retention interval conditions for subjects exposed to arbitrary forms as sample stimuli. Retention interval values are indicated above each graph.

The bottom graph shows that S5's latency to observing response was typically around 0.6 seconds during the first two conditions. The introduction of the two-second retention interval resulted in an increase in the latency to observing response, with median values around one second for this condition. During the four-second retention interval condition median latency to observing response increases sharply across sessions, peaking at over two seconds. This was the highest median latency to observing response for any subject in any condition. Latencies to observing response remained slightly high during the 8-second retention interval condition, but dropped to around one second for the remainder of the experiment.

Median latency to comparison selection for S5 (bottom graph) was consistently below one second during the zero and one-second retention interval conditions, then rose slightly during the 2-second condition. Latency to comparison selection increased across sessions within the RI=4" condition. From the RI=8" to RI=64" conditions, latency to comparison selection increased when retention interval was increased, although the effect was not as pronounced for S5 as it was for the other subjects.

In sum, latencies to observing response were generally short and consistent across conditions for all subjects, although S2, S4, and S5 showed slight increases across conditions. For S5, this increase in latency to produce comparison stimuli was transient. Latencies to observing response were longer for subjects exposed to tones as sample stimuli (S1, S2, and S3) than for subjects exposed to visual forms (S4 and S5). This is likely to be an artifact of the procedural requirement that the auditory sample play once completely before the program could recognize an observing response. For all subjects,

latency to comparison selection increased as a function of increases in retention interval value. Figure 8 shows latency to comparison selection plotted as a function of retention interval. Both axes have been scaled logarithmically to facilitate comparisons.

Conditions in which retention interval values were one second or more are shown.

Graphs in the left column show data from subjects S1, S2 and S3 top to bottom. Graphs for S4 and S5 appear in the upper right and middle right of the figure respectively. These five graphs also present the best fitting regression line (calculated by the least squares method). Presented in the bottom right graph are the regression lines from each individual subject as well as the regression line derived from the pooled data of all subjects. Taken as a whole, this figure further underscores the systematic changes in latencies to comparison selection as a function of changes in retention interval. In addition, the figure shows that S4 and S5's latencies to comparison were below 1 second in the first few conditions unlike latencies to comparison selection for S1, S2 and S3. Also, the straight lines in log-log coordinates suggest that the changes in latency to select comparison stimuli as a function of RI value would be well described by a power function.

Figures 9 through 13 show the results of error analyses for each subject. The top graphs of figures 9, 10, and 11 present trials in which the high-pitched tone was the sample and the letter "N" was the correct comparison. The middle graphs present trials where the middle-pitched tone was the sample and "M" was the correct comparison, and the bottom graphs show trials where the low-pitched tone was the sample and the correct comparison was "S". Figures 12 and 13 are similarly structured in terms of the comparison stimuli. The graphs in the left column of these figures show the cumulative

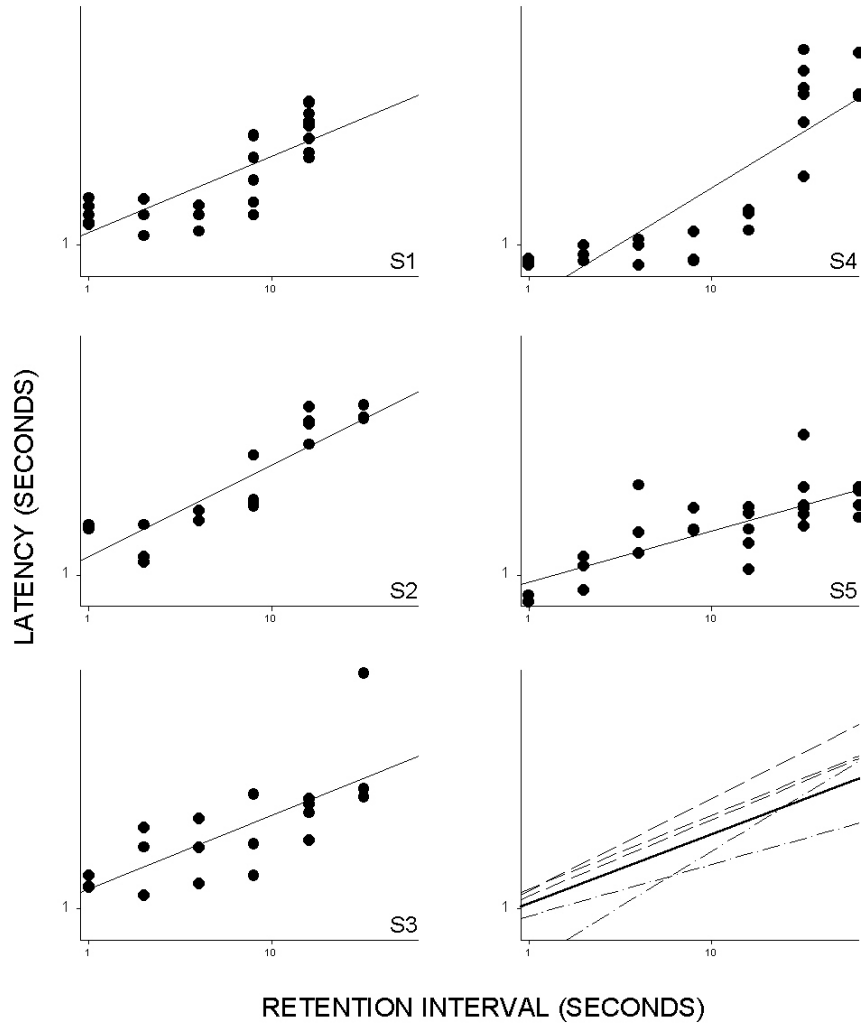


Figure 8. Median latency to comparison selection plotted as a function of retention interval. Each point represents a median value derived from a single session. Regression lines were fitted using the least squares method. The bottom right graph shows the individual regression lines from subjects exposed to tones as sample stimuli (dashed lines) and subjects exposed to arbitrary forms as sample stimuli (dash-dot) as well as a regression line fitted to the pooled data of all subjects (solid line).

number of errors by session. Graphs in the right column show the percentage of errors in each condition that were made to each of the two possible  $S^-$  comparison stimuli. For this reason the two bars (black and gray) representing a single condition will always sum to 100. The only exception to this is when no errors were made in a condition, in which case both bars are assigned a value of zero.

Figure 9 shows the type of errors made broken down by sample type for S1. For the next five figures, the graphs in the left column show cumulated frequencies of selection of each of the two distracter stimuli across conditions. The graphs in the right column express the error frequencies relative to total errors during the condition. This graph allows a more direct examination of the distribution of errors across conditions. The graphs in the left column show that S1 was more likely to make errors given the middle and high pitch tones as samples relative to the low pitch tone as sample. The individual cumulative records reveal some orderly results. For example, the divergence of the plots in the top left graph indicate that in the presence of the high-pitched tone, “M” was selected more often than “S”. The middle left graph shows that selecting “N” and “S” were approximately equally likely when the middle-pitched tone was the sample. The bottom left graph shows that on trials in which the low-pitched tone was the sample, “M” was more likely to be selected than “N”.

The graphs in the right column further show that when the sample was either the high or low-pitched tone the proportion of errors were tightly constrained to the selection of “M”. Taken together, the graphs in the right column show that as retention interval

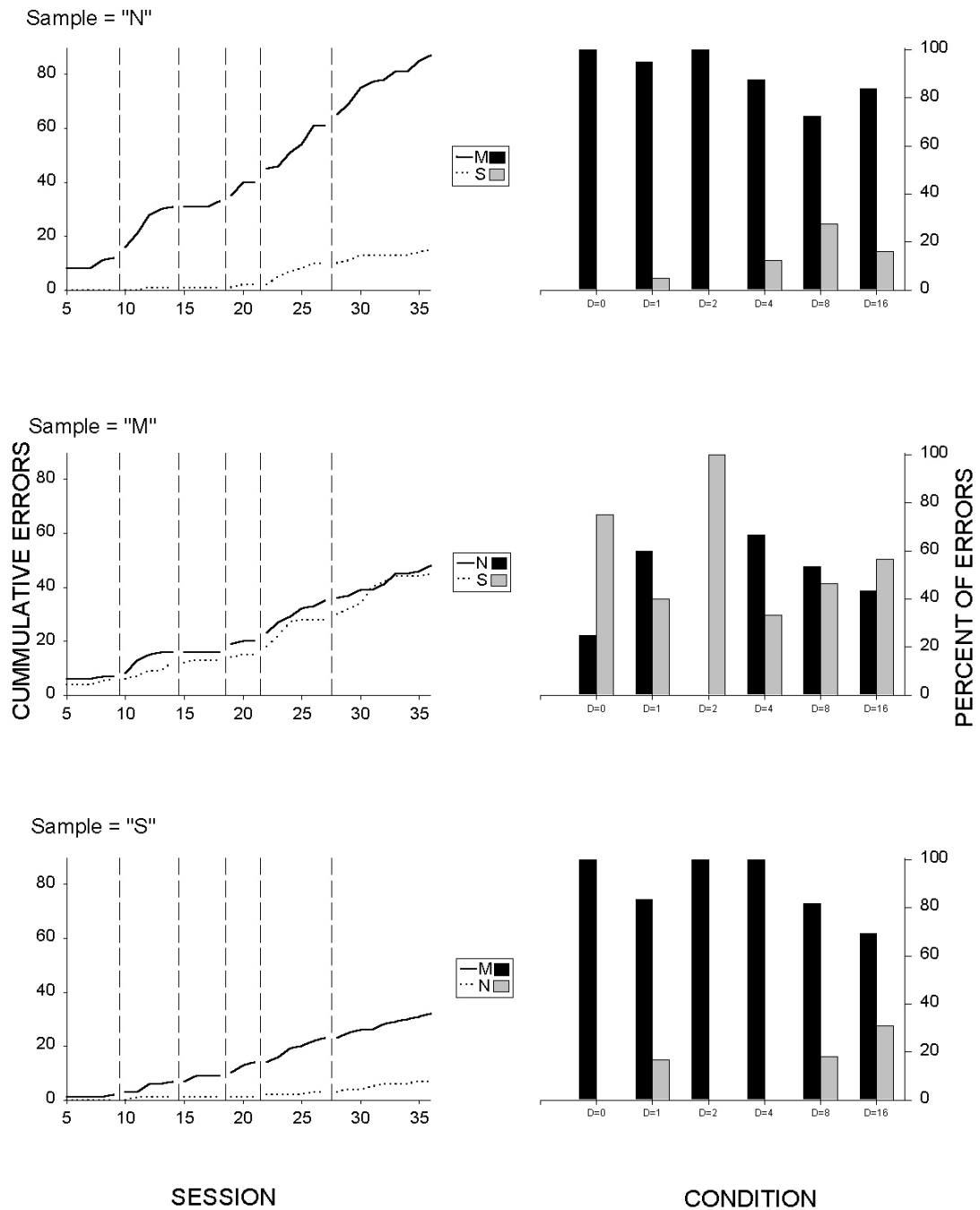


Figure 9. Error analysis for subject S1. Graphs in the left column show cumulative errors across retention interval conditions. Retention interval values left-to-right are 0, 1, 2, 4, 8, and 16 seconds. Graphs in the right column show percentage of error trials in which errors were made to either of two possible error types for a given sample, pooled by condition.



was increased the dispersal of errors widened, with errors becoming more evenly dispersed during the eight and 16-second retention interval conditions.

Figure 10 presents error analysis for subject S2. The cumulative error graphs in the left column show that errors occurred at similar rates on trials for which high and middle-pitched tones were samples, and that the least errors occurred when the low-pitched tone was the sample. The top and bottom left graphs also show that there was a substantial over-selection of “M”. When the middle tone was the sample errors were more dispersed, although there was a substantial over-selection of “N”. The graphs in the right column show that there is a slight increase in the dispersal of errors as retention interval is increased, but that when retention interval was two seconds or less error trials in which the sample was either a high or low-pitched tone were marked by exclusive selection of “M”.

Figure 11 shows error analysis for S3. Graphs in the left column show that compared to S1 and S2, subject S3 made relatively few errors. Errors were slightly more likely to occur on trials in which the high-pitched tone was the sample than when the middle or low-pitched tones were samples. An equal number of errors occurred when the samples were the middle and low-pitched tones. Left column graphs show over-selection of “M” when a high or low-pitched tone was the sample. The bottom right graph shows that when the low-pitched tone was the sample errors were made exclusively to “M” throughout all delayed-matching-to-sample conditions. The top right graph, however, shows increased dispersal of errors when retention interval is 16 or 32 seconds.

In sum, figures 9, 10, and 11 show that for subjects exposed to tones as sample stimuli errors were more likely to occur on trials with a higher frequency sample than on

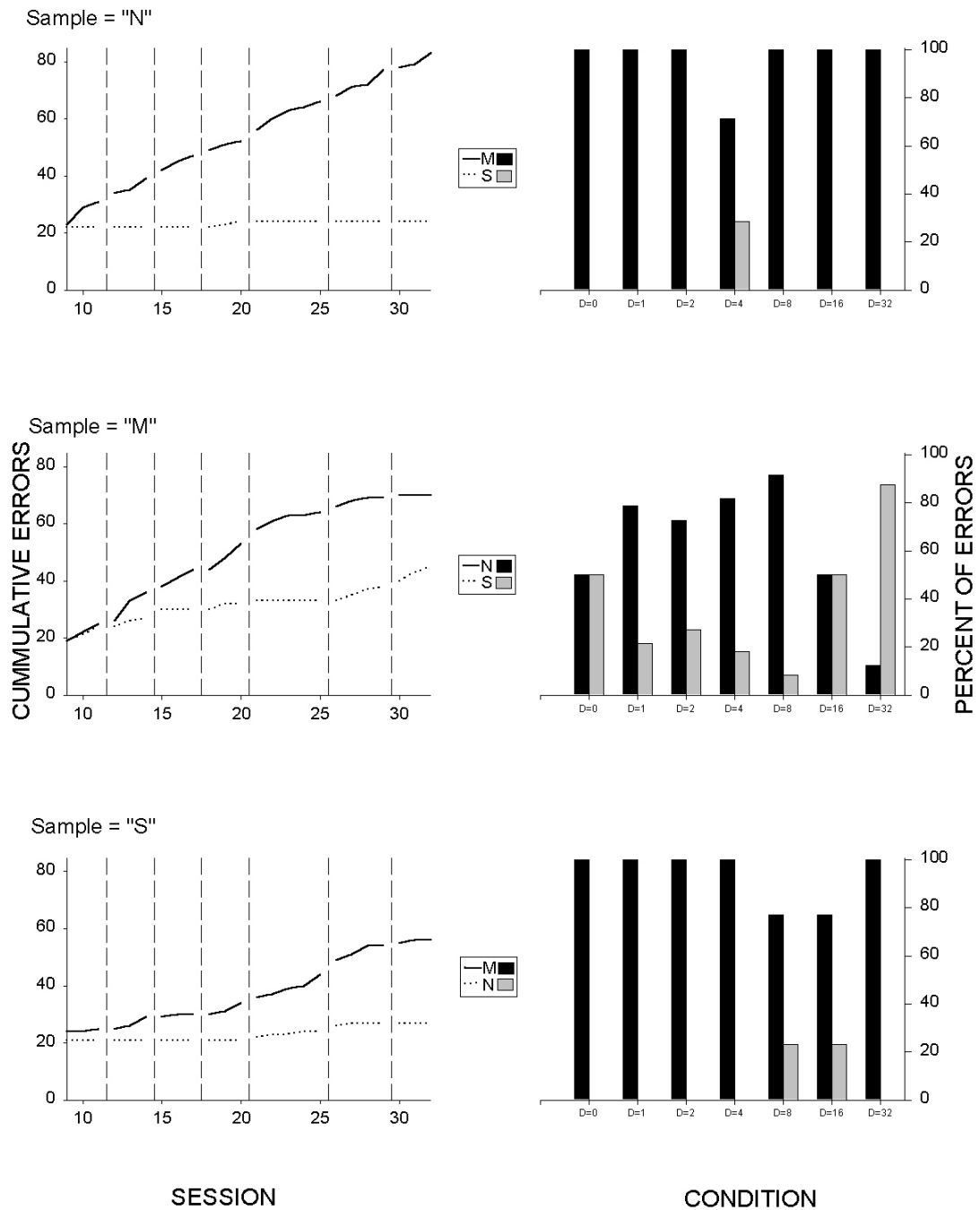


Figure 10. Error analysis for subject S2. Graphs in the left column show cumulative errors across retention interval conditions. Retention interval values left-to-right are 0, 1, 2, 4, 8, 16, and 32 seconds. Graphs in the right column show percentage of error trials in which errors were made to either of two possible error types for a given sample, pooled by condition. Note that the scale of the Y-axis in the graphs on the right differs from the scale of similar error analysis graphs presented in the prior figure.

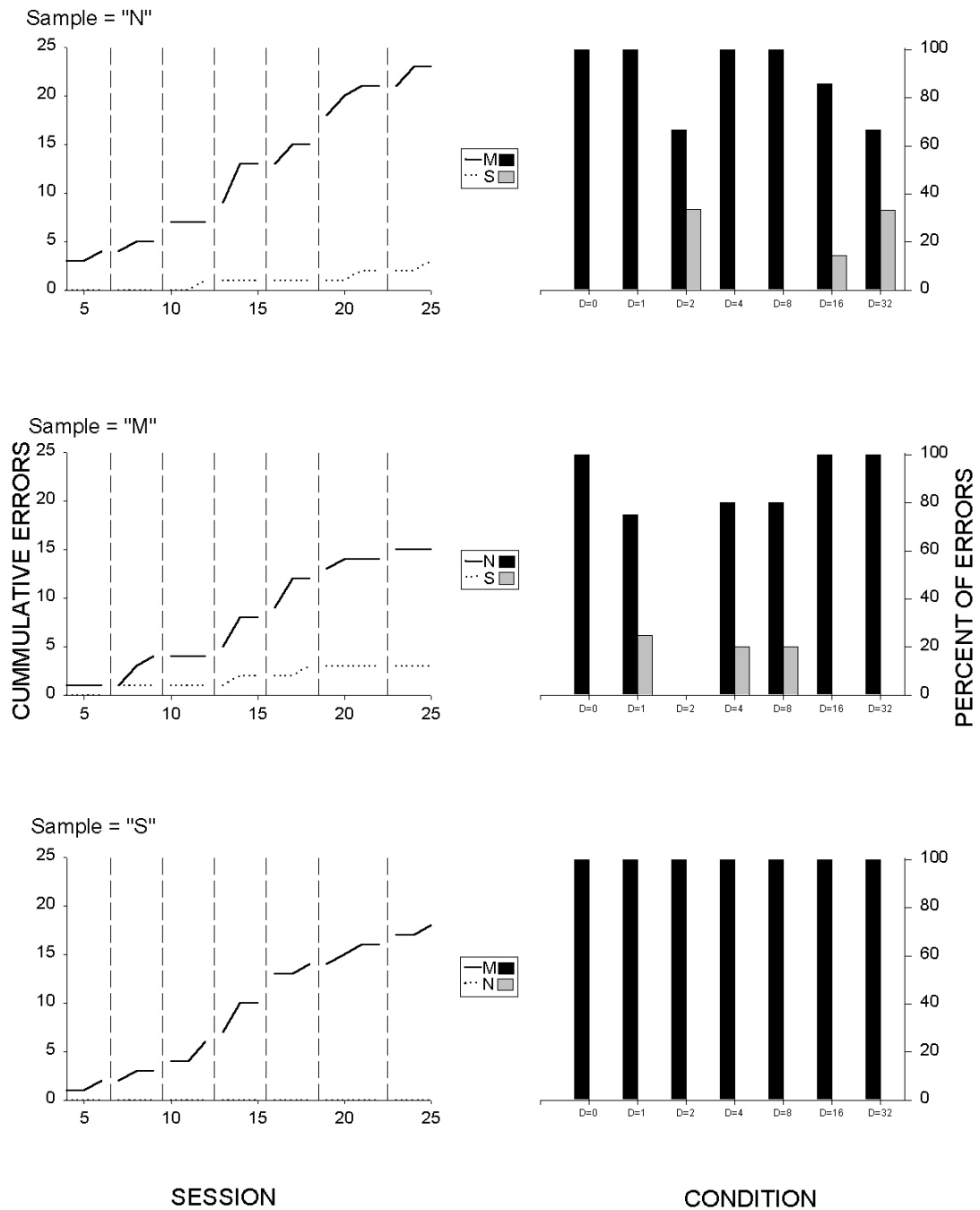


Figure 11. Error analysis for subject S3. Graphs in the left column show cumulative errors across retention interval conditions. Retention interval values left-to-right are 0, 1, 2, 4, 8, 16, and 32 seconds. Graphs in the right column show percentage of error trials in which errors were made to either of two possible error types for a given sample, pooled by condition. Note that the scale of the Y-axis in the graphs on the right differs from the scale of previous similar error analysis graphs.

trials with a lower frequency sample. On error trials in which the sample was the high pitch tone ( $N=S^+$ ) or low pitch tone ( $S=S^+$ ) subjects were more likely to pick the comparison designated as the  $S^+$  on trials in which the middle pitch tone was the sample (M). On error trials in which the sample was the middle pitch ( $M=S^+$ ) two of three subjects (S2 and S3) selected the comparison designated as the  $S^+$  on trials in which the high frequency tone was the sample (N), and one subject (S1) was equally likely to select either of the incorrect comparison stimuli. This pattern of error dispersal however, was not as pronounced as that of the over-selection of the “M” comparison stimulus. Lastly, as retention interval lengthened all subjects showed a slightly more even dispersal of errors.

Figure 12 shows error analysis for S4. Graphs in the left column show an extremely low overall occurrence of errors. Graphs in the right column show no systematic change in dispersal of errors as a function of increases in retention interval. Taken together, these six graphs show no systematic pattern of errors.

Figure 13 presents analysis of errors for subject S5. The cumulative error graphs in the left column show that errors occurred more frequently for S5 than for S4, but that the overall number of errors was still low. The top and middle graphs of the left column show over-selection of the comparison stimulus “S” in the presence of samples for which “N” and “M” was the  $S^+$ . The bottom left graph shows that errors were less likely when the sample stimulus is the form for which “S” was the  $S^+$ , and that when errors are made on such trials the comparison stimuli “N” and “M” were equally likely to be selected. Graphs in the right column show no increase in dispersal of errors as a function of

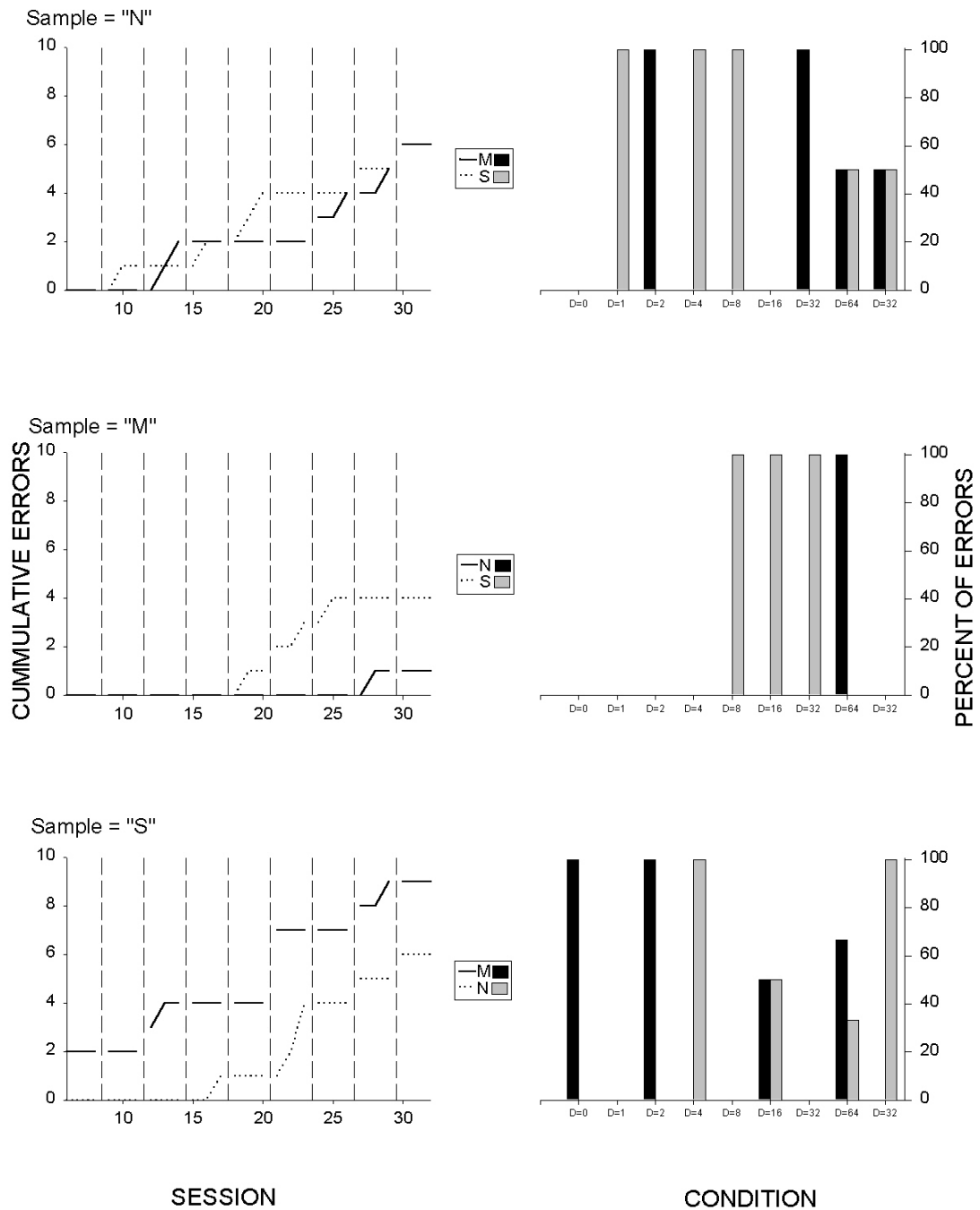


Figure 12. Error analysis for subject S4. Graphs in the left column show cumulative errors across retention interval conditions. Retention interval values left-to-right are 0, 1, 2, 4, 8, 32, 64 and 32 seconds. Graphs in the right column show percentage of error trials in which errors were made to either of two possible error types for a given sample, pooled by condition. Note that the scale of the Y-axis in the graphs on the right differs from the scale of previous similar error analysis graphs.

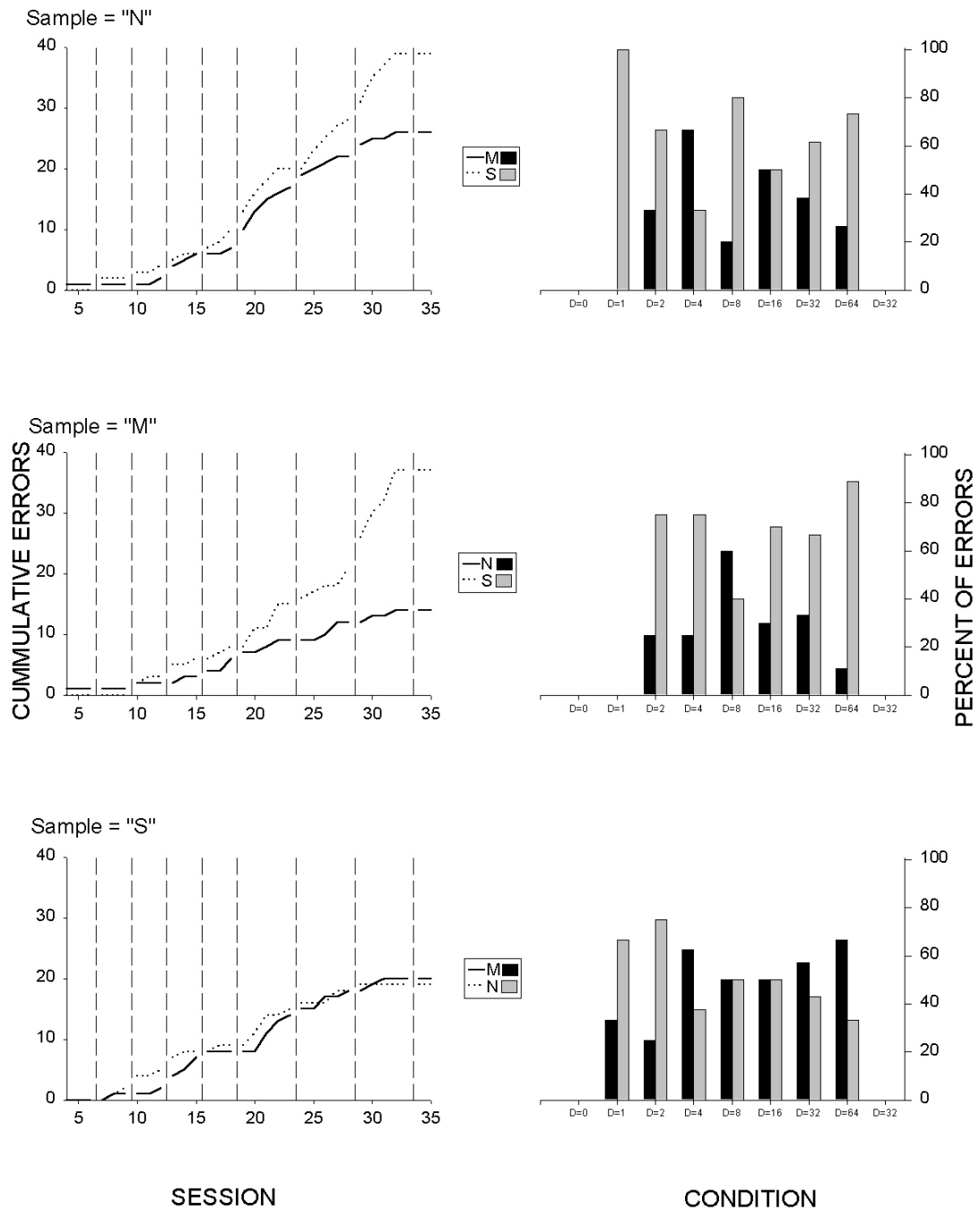


Figure 13. Error analysis for subject S5. Graphs in the left column show cumulative errors across retention interval conditions. Retention interval values left-to-right are 0, 1, 2, 4, 8, 32, 64 and 32 seconds. Graphs in the right column show percentage of error trials in which errors were made to either of two possible error types for a given sample, pooled by condition. Note that the scale of the Y-axis in the graphs on the right differs from the scale of previous similar error analysis graphs.

increases in retention interval. When retention interval is changed from 32 seconds to 64 seconds the bias toward over-selection of “S” when “N” or “M” are the  $S^+$  is increased slightly.

In sum, figures 12 and 13 show that subjects exposed to visual forms as sample stimuli generally made few errors. The systematic patterns of errors shown in figures 9, 10, and 11 for subjects S1, S2, and S3, respectively, were not repeated in figures 12 and 13 for subjects S4 and S5. Furthermore, increases in retention interval resulted in increases in error distribution across conditions for subjects S1, S2, and S3, but not for subjects S4 and S5.

## CHAPTER 5

### DISTRACTION

#### Procedure

A procedure commonly used to increase the difficulty involves adding putatively distracting response requirements during the retention interval (e.g., Torgrud and Holborn, 1989). This procedure was employed in the current study to investigate the possibility that accuracy was relatively stable across retention intervals because subjects were using a differential response (such as naming and rehearsing the sample stimuli) to bridge the delay between sample offset and comparison stimulus onset. Specifically, this phase of the experiment served as an attempt to disrupt stimulus control by requiring a variety of tasks during the retention interval. All subjects were given the following instruction before their first distraction condition:

Things may be slightly different today but remember that all responses are to be made by touching the screen.

Four different distraction conditions were programmed for S4. The remainder of the subjects were exposed to only two of the four distraction conditions. One distraction condition exposed the subject to a series of nine simultaneous conditional color-matching trials during the retention interval. The precise length of the retention interval was determined by the amount of time required to complete the color matching trials, but the number of color matching trials was programmed to approximate 32 seconds. Another distraction condition (sample distraction) exposed subjects to 12 presentations of stimuli used as samples in quasi-random order and position on the screen. Stimuli were



presented singly and subjects were required to touch each stimulus or its visual accompaniment to advance to the next stimulus. The length of the retention interval was again determined by the amount of time required to complete the task, but approximated the retention interval experienced by each subject in the previous condition. A similar distraction condition (comparison distraction) required subjects to touch stimuli used as comparisons in identical fashion to the sample distraction condition. Subjects were also exposed to a condition (reading distraction) in which they were required to read out loud during the retention interval. This condition was different from all previous conditions in that the experimenter sat in the laboratory room with the subject while the session was being conducted. Subjects were given the following instruction before their first reading distraction session:

Today you will be required to read during your session. Please pick one of these books to read. I will sit in the room with you while you read out loud. When it is time to make your response I will say “ready”.

During the retention interval the screen was entirely white. The subject read out loud from one of four books of short stories provided by the experimenter, which was chosen by the subject before beginning each reading distraction session. Subjects were not permitted to read the same story more than once, but no other restrictions were placed on their selection of reading materials. When the comparison stimuli appeared the experimenter said, “ready” and the subject touched the screen as in all other conditions. The retention interval was set to equal the last retention interval experienced by each subject in the delayed-matching-to-sample phase.

Subject S4 was exposed to the general identity matching (GIM), sample (SAM), comparison (COM), and reading (REA) distraction conditions, in that order. All other subjects were exposed to the sample distraction condition first, followed by the reading distraction condition. S3 was returned to the sample condition for a single session. Due to time constraints stability requirements for changing conditions were occasionally waived.

## Results

Figure 14 presents the percentage of correct trials for all subjects across all distraction conditions. Data from the last delayed-matching-to-sample condition experienced by the subject prior to introduction of the distraction condition are presented again to facilitate comparison. Graphs in the left column show data for subjects S1, S2, and S3, respectively. For S1, the top left graph shows an initial decrease in percentage of correct trials during the first sample distraction session, followed by a substantial recovery in the second session. After the sample distraction condition a reading distraction condition was introduced. Percentage of correct trials during the reading sessions was consistently around 85%, which is within the range of percent of correct trials in the sample distraction condition. This is a decrease in accuracy compared to the 16-second retention interval condition. For S2, the middle left graph shows that accuracy increased during the first sample distraction condition compared with the 32-second retention interval. The reading distraction condition was then introduced. S2's performance was less accurate on three of the four reading distraction conditions compared to the 32-second interval delayed-matching-to-sample condition. For S3, the

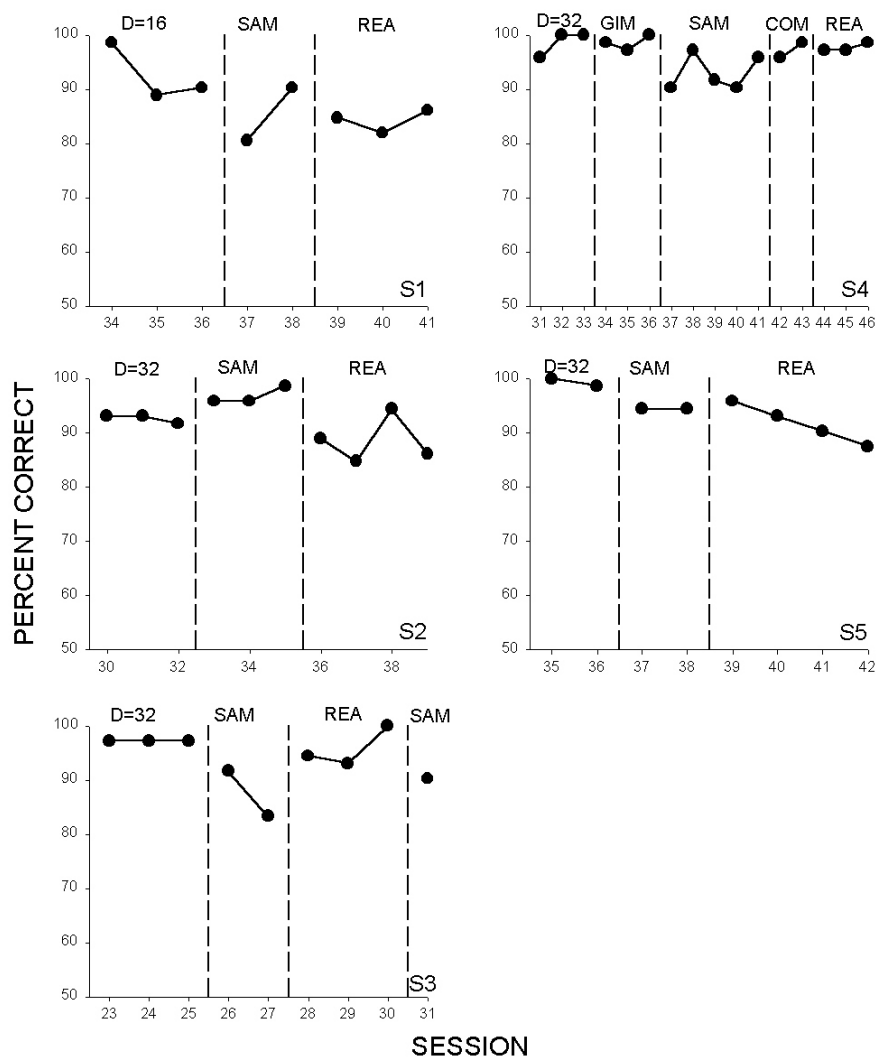


Figure 14. Percentage of correct responses across distraction conditions. The following labels are assigned to each condition; generalized identity matching distraction condition = GIM, sample distraction condition = SAM, comparison distraction condition = COM, reading distraction condition = REA. Data from the final DMTS condition are re-presented to facilitate comparison.

bottom left graph shows a substantial decrease in accuracy during the sample distraction condition. S3's accuracy during the reading distraction condition was higher than accuracy during the sample distraction condition, but lower than accuracy during the 32-second retention interval condition.

Graphs in the right column show data from subjects S4 and S5. For S4, the top graph shows that the accuracy of performance was not affected during the general identity matching distraction condition. Percent of correct trials decreased during the sample distraction condition, but recovered during both the comparison distraction and reading distraction conditions. The bottom right graph (S5) shows that accuracy decreased slightly during the sample distraction condition. Performance improved slightly during the first exposure to the reading distraction condition but became less accurate within the condition, with a clear downward trend over the four sessions of this condition.

In sum, all five subjects show a decrease in percent of correct trials during at least one distraction condition when compared to performance on the final delayed-matching-to-sample condition. Only S2 did not show at least a slight decrease in accuracy during the sample distraction condition, and three of five subjects (S1, S2, and S5) showed less accurate performance during the reading distraction condition than during conditions with identical retention intervals in which there were no formal response requirements. The identity matching and comparison distraction conditions did not affect percent of correct trials for the only subject exposed to those conditions (S4).

Error analyses for all subjects across all distraction conditions are presented in Table 2. One finding from previous error analysis concerning dispersal of errors for subjects exposed to tones as sample stimuli was apparent in distraction conditions as well. On error trials in which the sample was the high pitch tone ( $N=S^+$ ) or low pitch tone ( $S=S^+$ ) there was a relatively narrow dispersal of errors, with subjects picking the comparison designated as the  $S^+$  on trials in which the middle pitch tone was the sample (M). The systematic dispersal of errors, apparent in the performance of subjects exposed to tones as sample stimuli, is the only consistent pattern of error during distraction conditions for any of the five subjects.

After the completion of the distraction conditions subjects were debriefed. During these debriefings all subjects reported the use of differential coding responses. For all subjects the reported response topography was identical. Specifically, all subjects reported saying the name of a letter being used as a comparison stimulus immediately upon exposure to the sample stimulus. Furthermore, subjects reported engaging in precurrent behavior during the retention intervals. Four of the five subjects performed some sort of rehearsal, and one subject (S2) positioned his index finger on the computer desk differentially dependent upon the sample.

Subject	Condition	Session #	S <sup>+</sup> = N		S <sup>+</sup> = M		S <sup>+</sup> = S	
			M	S	N	S	M	N
S1	SAM	37	3	0	1	3	3	1
		38	1	1	2	1	1	1
	REA	39	4	2	0	0	5	0
		40	4	0	1	2	3	3
		41	1	4	1	4	0	0
S2	SAM	33	2	0	1	0	0	0
		34	1	0	2	0	0	0
		35	1	0	0	0	0	0
	REA	36	3	0	2	1	2	0
		37	4	0	2	1	3	1
		38	3	0	1	0	0	0
		39	7	0	0	2	1	0
S3	SAM	25	2	1	2	0	1	0
		26	1	1	1	4	2	3
	REA	27	2	0	0	1	0	0
		28	2	0	0	0	3	0
		29	0	0	0	0	0	0
	SAM	30	0	1	1	0	1	1
S4	GIM	33	1	0	0	0	0	0
		34	1	0	0	1	0	0
		35	0	0	0	0	0	0
	SAM	36	2	1	0	2	0	0
		37	0	1	0	0	1	0
		38	3	1	1	0	0	1
		39	3	0	0	0	1	3
		40	1	1	0	0	0	1
	COM	41	2	0	0	0	0	0
		42	1	0	0	0	0	0
	REA	43	0	0	0	0	1	1
		44	1	0	0	0	1	0
		45	0	0	0	0	0	1
S5	SAM	36	2	0	1	0	1	0
		37	2	1	0	0	0	1
	REA	38	0	0	1	1	1	0
		39	0	2	1	0	1	1
		40	2	1	0	3	0	1
		41	1	4	3	0	0	1

Table 2. Error analysis for distraction conditions. The numbers in each cell represent the number of times in that session in which a subject chose a particular comparison stimulus in the presence of a particular sample stimulus.

## CHAPTER 6

### DISCUSSION

This experiment compared the performances of human subjects on arbitrary conditional discrimination tasks. Two subjects were exposed to a conditional discrimination task in which arbitrary visual forms served as sample stimuli and upper-case English letters served as comparison stimuli. For the other three subjects, auditory stimuli (pure tones that were 1-second in duration) were used as sample stimuli with upper-case English letters serving as comparison stimuli. The data show that overall accuracy decreased and latencies to select comparison stimuli increased as the value of the retention interval was increased. The data show a further decrease in accuracy when response requirements were imposed during the retention interval. Finally, there were small but systematic differences in the nature of the errors made by subjects exposed to visual versus auditory stimuli. Each of these findings will be discussed.

For all five subjects, latency to comparison selection increased as retention interval increased. Lines fitted to the distribution of scores appear to be described by a power function (see figure 8). Baron and Menich (1985) also found that although errors did not change systematically as retention interval was varied, increases in retention interval resulted in slower responding to comparison stimuli. Taken together, these findings may suggest that latency and accuracy measures are dissociable and changes in their obtained values may result from different mechanisms. This should inform our discussion of latency as a measure of stimulus control. If latency and accuracy measures consistently co-varied it would be reasonable to address stimulus control questions based

solely on one or the other of these measures. However, the dissociation of these measures leaves open several possibilities. For example, latency may be a more sensitive measure of stimulus control. Alternatively, that accuracy and latency may reflect independent or at least separable aspects of stimulus control. Conclusions about which of these possibilities, if any, are tenable, cannot be drawn from the present data set.

As stated previously, a decrease in accuracy as a function of increases in retention interval is the most common finding in delayed matching-to-sample preparations in which retention interval is an independent variable. The methods used in the present experiment, however, failed to produce the robust effects common in similar studies of non-human DMTS performance. There are several factors that may account for this discrepancy.

One possible explanation is that the retention intervals may have been too brief to affect human performance on a DMTS task. After extensive training at zero delay conditions, bottlenose dolphins showed perfect accuracy on DMTS tasks with retention intervals as high as 120 seconds (Herman & Gordon, 1974). It is not entirely surprising then, that humans showed accurate performance with retention intervals as long as 64 seconds.

Although retention intervals in the present experiment may have been too brief, there were enough errors made to enable a revealing analysis. This analysis warrants a brief diversion from the discussion of why accuracy remained relatively high across retention intervals. Unlike accuracy measures, which do not indicate a clear difference between subjects, the error analyses do show a clear difference in the dispersal of the errors made by subjects exposed to auditory stimuli as samples compared to subjects



exposed to visual stimuli. Error analyses show that subjects exposed to auditory sample stimuli were considerably more likely to make errors to comparison choices that were closer to the correct choice in terms of frequency of the sample stimulus. For example, if the sample was either the high or low frequency tone, errors were most likely to occur to the comparison stimulus that served as the  $S^+$  on trials in which the middle frequency tone was the sample. Subjects exposed to visual sample stimuli showed no such organization of errors. The difference is likely due to the fact that the auditory stimuli are organized along a single stimulus dimension (namely, the frequency of the wavelength). Sidman (1969) showed standard generalization gradients on delayed matching tasks in which the sample stimuli were ellipses that varied along the dimension of height. The errors of subjects exposed to ellipses as sample stimuli in Sidman's experiment were organized similarly to the errors of subjects exposed to auditory sample stimuli in the present study. Furthermore, the finding that errors become more widely dispersed as retention interval increased in the present study is similar to Sidman's finding that generalization gradients flattened as retention interval was increased. Considering the present findings in the context of Sidman's data, we cannot conclude that the difference in the nature of errors between subjects exposed to audio versus visual sample stimuli was a function of sample stimulus modality. Further investigation is required to answer questions concerning the organization of errors based on properties of the sample stimuli. Particular investigations are suggested below.

Other properties of the sample stimuli also may have affected the accuracy measures in the present experiment. First, sample stimulus discriminability is a critical determinant of accuracy on DMTS tasks (Carter & Eckerman, 1975). The arbitrary

forms used as sample stimuli for subjects exposed to visual samples did not vary along a common dimension, nor did they share any obvious common features. Results from S4 and S5 indicate that this may have resulted in a high level of sample stimulus discriminability. More recently, duration of the sample stimulus has been shown to play a role in performance on DMTS tasks. Specifically, increased exposure to sample stimuli increases initial discriminability between stimuli (White, 1985). The sample duration used for auditory stimuli in this study (one second) may have been a critical factor in the overall high level of initial discriminability shown by all subjects exposed to auditory sample stimuli (see  $D=0$  in figure 3).

Second, comparison stimulus discriminability has also been shown to influence accuracy on DMTS tasks (Carter & Eckerman, 1975). In considering the experimental subjects (undergraduate students) the experimenters chose upper-case English letters as comparison stimuli explicitly to facilitate a high level of discriminability between comparison stimuli. The high level of comparison discriminability and familiarity is a likely factor in the overall high accuracy across subjects. Furthermore, all subjects in the present study reported during debriefing that they made naming responses during the experimental sessions. The subjects' familiarity with comparison stimuli, along with the fact that the stimuli have commonly used "names" may have made it easier for subjects to name the stimuli. Generally these naming responses took the form of rehearsing the name of a comparison letter. For instance, S3 reported that on trials in which the low frequency tone was the sample, he would covertly repeat the letter "S" to himself. As a function of previous reports of this phenomenon, this finding was anticipated.

Distraction conditions were constructed specifically to interfere with such rehearsal. The

fact that accuracy decreased slightly during at least one distraction condition for all subjects indicates that these procedures were mildly effective in preventing rehearsal. During debriefing subjects reported that although it was somewhat more difficult to do so, they continued to rehearse during the retention interval.

Lastly, that performance frequently improved within conditions indicates that subjects may have been learning with each new condition. Sargisson & White (2001) found that retention interval during training greatly influenced performance during testing at a series of retention intervals. That is, performance was more accurate at longer retention intervals when those intervals matched the retention interval experienced during training of the conditional relations. These authors concluded that the retention interval itself entered into the controlling relations determining performance on the DMTS task. The present experiment may be conceived of as training subjects to respond at each retention interval because of the procedural requirement that subjects were exposed to a single retention interval within a session and that subjects were exposed to the same retention interval across sessions until accuracy stabilized.

Ultimately, this preparation has not produced a definitive answer to our original questions. Some procedural changes will be required if stronger conclusions about the influence of stimulus modality on human performance in DMTS tasks are to be made. Changes such as decreasing sample duration for auditory sample stimuli and increasing retention intervals could potentially contribute to a more sensitive preparation for the present experimental question. Furthermore, exposing subjects to all retention intervals within each session may prevent the retention interval from entering into the controlling relations, further disrupting performance, especially at the longer retention intervals.

The results of this experiment have generated a host of new experimental questions and directions of future research. First, a direct comparison of auditory and visual stimuli in which visual sample stimuli differed from one another along a single common dimension would allow for a more systematic analysis of the organization of errors in a delayed matching task. This same preparation may be used to yield information about cross-modal differences in discriminability between sample stimuli by comparing properties of visual stimuli (e.g., ellipses of varying height, colors of varying wavelength) to properties of audio stimuli (e.g., frequency, amplitude). Another avenue for future research is a more in-depth investigation of features of comparison stimuli in conditional discriminations. When compared to the role of sample stimuli, little is known about the role of comparison stimuli in DMTS performance. Studies manipulating the familiarity, complexity and sample-to-sample differences along a common dimension would all aid in elucidating the role of comparison stimulus features in investigations of remembering. It is possible that the context in which one remembers is critical to the behavior of remembering, and the comparison array is clearly an important part of that context in a DMTS task.

In sum, although the data presented here do not allow for an answer to the original experimental question, they are useful in a variety of ways. First, the small, yet systematic, effects suggest that the independent variables chosen in this experiment do indeed play a role in remembering, but that the parameters of these variables must be adjusted to answer similar questions in the future. Second, these results, when taken together with the existing literature, suggest specific procedural changes (e.g. decreased sample duration, increased retention intervals) that would provide more sensitive

preparations for questions concerning the phenomenon of remembering. Third, this experiment serves as an impetus for future research concerning the interaction between the original sample stimulus and the context in which remembering occurs.

## REFERENCES

- Baron, A., & Menich, S. R., (1985). Reaction times of younger and older men: Effects of compound samples and a prechoice signal on delayed matching-to-sample performances. Journal of the Experimental Analysis of Behavior, 44(1), 1-14.
- Blough, D. S., (1959). Delayed matching in the pigeon. Journal of the Experimental Analysis of Behavior, 2(X), 151-160.
- Carter, D. E., & Eckerman, D. A. (1975). Symbolic matching by pigeons: Rate of learning complex discriminations predicted from simple discriminations. Science, 187(4177), 662-664.
- Critchfield, T. S., & Perone, M., (1990). Verbal self-reports of delayed matching to sample by humans. Journal of the Experimental Analysis of Behavior, 53(3), 231-344.
- Dube, W. V., Rocco, F. J., & McIlvane, W. J., (1989). Delayed matching to sample with outcome-specific contingencies in mentally retarded humans. Psychological Record, 39(X), 483-492.
- Herman, L. M., & Gordon, J. A., (1974). Auditory delayed matching in the bottlenose dolphin. Journal of the Experimental Analysis of Behavior, 21(1), 19-26.
- Hunt, M., Parr, W. V., & Smith, P., (1999). Local and global sources of control in pigeon delayed matching-to-sample performance. Quarterly Journal of Experimental Psychology, 52B(3), 203-233.
- Jones, B. M., & White, K. G., (1992). Sample stimulus discriminability and sensitivity to reinforcement in delayed matching to sample. Journal of the Experimental Analysis of Behavior, 58(1), 159-172.
- Kraemer, P. J., & Roberts, W. A., (1984). Short-term memory for visual and auditory stimuli in pigeons. Animal Learning and Behavior, 12(3), 275-284.
- Nelson, K. N., & Wasserman, E. A., (1978). Temporal factors influencing the pigeon's successive matching-to-sample performance: Sample duration, intertrial interval, and retention interval. Journal of the Experimental Analysis of Behavior, 30(2), 153-162.
- Palmer, D. C., (1991). A behavioral interpretation of memory. In L. J. Hayes & P. N. Chase (Eds). Dialogues on verbal behavior: The first international institute on verbal relations, Reno, NV: Context Press (261-279).

Parr, W. V., Hunt, M., & Williams, W. A., (1999). Sample frequency and sample duration as sources of stimulus control in delayed matching to sample. Behavioral Processes, 46(1), 39-55.

Parsons, J. A., Taylor, D. C., & Joyce, T. M., (1981). Precurrent self-prompting operants in children: "Remembering". Journal of the Experimental Analysis of Behavior, 36(2), 253-266.

Roediger, H. L., Marsh, E. J., & Lee, S. C., (2002). Kinds of memory. In H. Pashler and D. Medin (Eds). Stevens' Handbook of Experimental Psychology (3<sup>rd</sup> Edition) Volume 2: Memory and Cognitive Processes, John Wiley & Sons, Inc. (1-42).

Sargisson, R. J., & White, K. G., (2001). Generalization of delayed matching to sample following training at different delays. Journal of the Experimental Analysis of Behavior, 75(1), 1-14.

Schneider, W., (2000). Research on memory development: Historical trends and current themes. International Journal of Behavioral Development, 24(4), 407-420.

Sidman, M., (1969). Generalization gradients and stimulus control in delayed matching-to-sample. Journal of the Experimental Analysis of Behavior, 12(5), 745-757.

Torgrud, L. J., & Holborn, S. W., (1989). Effectiveness and persistence of precurent mediating behavior in delayed matching to sample and oddity matching with children. Journal of the Experimental Analysis of Behavior, 52(2), 181-191.

White, K. G., (1985). Characteristics of forgetting functions in delayed matching to sample. Journal of the Experimental Analysis of Behavior, 44(1), 15-34.

Wixted, J. T., & Ebbesen, E. B., On the form of forgetting. Psychological Science, 2(6), 409-415.